ML Assignment 1

Neeharika Immaneni - NI2452

neeharika@utexas.edu

1 Introduction

The goal of the assignment was to use independent component analysis (ICA) for accomplishing the Blind Source problem (also referred to as the cocktail party problem) which involves the separation of a set of source signals from mixed signals, without any information about the source signals or the mixing process. We are provided few source signals which are mixed and these mixed signals are processed to recover the original signals. Maximum accuracy is obtained if the source signals are minimally correlated and are completely independent of each other. Many experiments were done which have been elaborated in detail below.

2 Method

The sound signals, which are non-guassian are loaded into vectors in MATLAB. If the input sources are guassian in nature, they would be rotationally symmetric and it is not possible to distinguish them using ICA i.e. we assume the input vector is not multi variate Guassian.

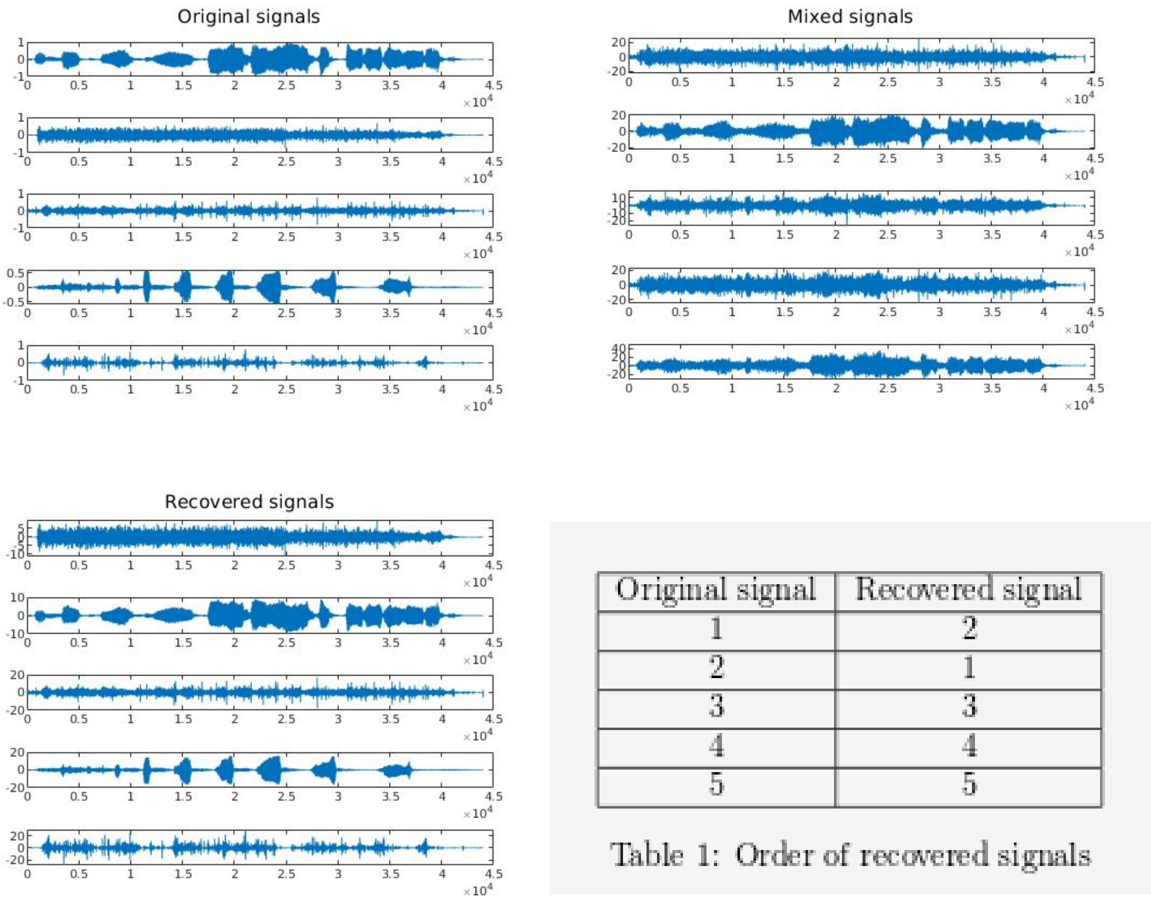
A mixing matrix (X) is created by multiplying the source signals(U) each of length t, with a matrix comprising of random real numbers (A), where Ai,j is the weight of the jth source signal in the ith mixed signal.

$$X = A * U \tag{1}$$

The method involves determining a matrix W, that helps us to recover the original signals. We need to find out a method of determining the log-likelihood of the input signals that depends on the parameter W, and then perform updates iteratively to improve our estimated W (which is how stochastic gradient descent works). Until convergence, we keep computing W, to get the output signals Y. The below steps are repeated till it converges. Here Z is a sigmoid function and using the stochastic gradient descent (SGD) update, we compute the value of W thereby moving in the direction of the gradient descent.

$$\begin{aligned}
&Y = W * X \\
&Z = 1/(1 + exp(-Y)) \\
&delta_W = learningrate * (I + (1 - 2Z)Y^T/t) * W \\
&W = W + delta_W
\end{aligned} \tag{2}$$

To avoid exploding of the W matrix, division is performed by the signal length.



Original signal	Recovered signal
1	2
2	1
3	3
4	4
5	5

Table 1: Order of recovered signals

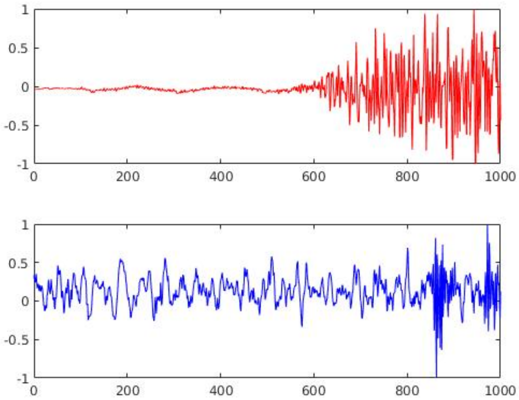
Figure 1: Signal Reconstruction

3 Experiments

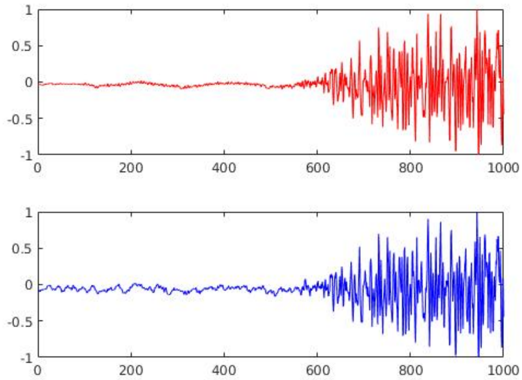
Various experiments were performed by varying the learning rate and the number of iterations. The original, mixed and reconstructed signals are shown in Figure 1 with a learning rate of 0.01 and 100000 iterations. The mixing was performed by varying the number of input signals (1 to 5), since they were all completely not correlated of each other (e.g. a man speaking, a person's laugh etc.) even while considering all the 5 signals, recovery of the individual sources was possible. The task could have been more challenging, given speech input of different people of the same gender conversing.

3.1 Varying the learning rate

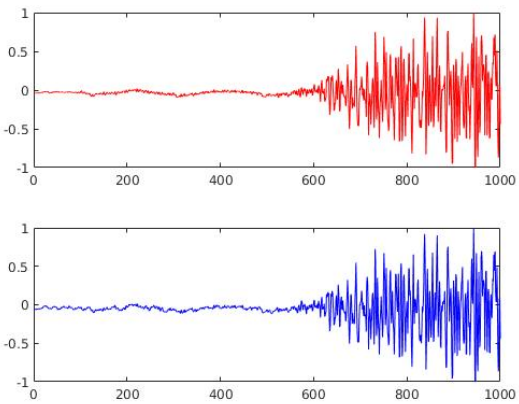
Experiments were performed on all the signals, for 1000 iterations by varying the learning rate and it can be observed as shown in Figure 2 that as the learning rate increases, the accuracy also increases since the convergence of the gradient descent is better. The input and corresponding recovered signal for the first 1000 time periods for one of the sound signals is mapped. Initially when the learning rate is very low as 0.001, the output is completely different from the input signal and the correlation was approximately 0.6. With a learning rate of 0.01, for the initial 600 time frames, the model still wasn't able to reproduce the original signal, however for learning rate of ¿0.05 the reconstructed signal starts to look like the original signal with the correlation value of 0.99.



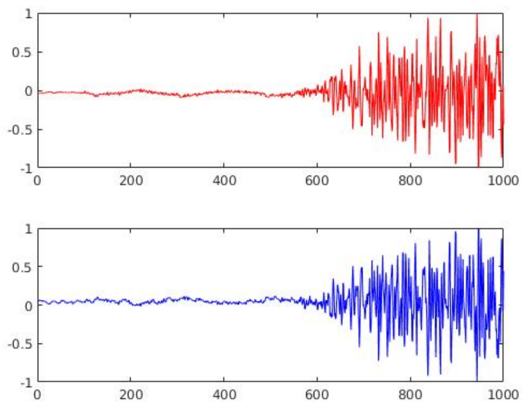
(a) Learning rate = 0.001



(b) Learning rate = 0.01



(c) Learning rate = 0.05



(d) Learning rate = 0.09

Figure 2: Reconstruction of signals with varied learning rates (RED - INPUT, BLUE - OUTPUT)

3.2 Accuracy

Since this is a blind recovery problem, the order of the recovered signals is different and I used the correlation co-efficient to compare each input signal, with all the recovered output signals and the one with the highest correlation is mapped as it's corresponding input signal.

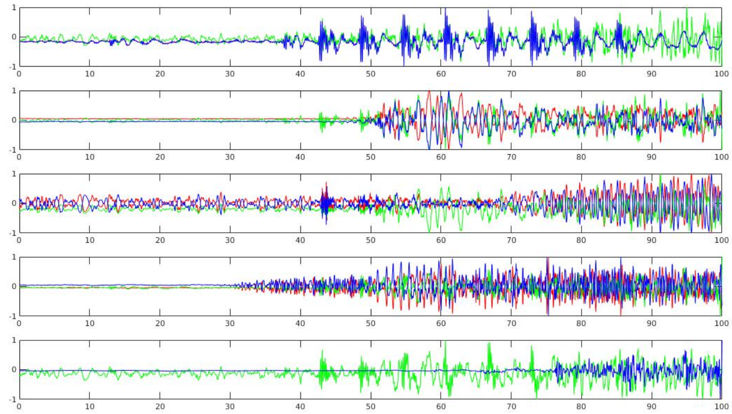


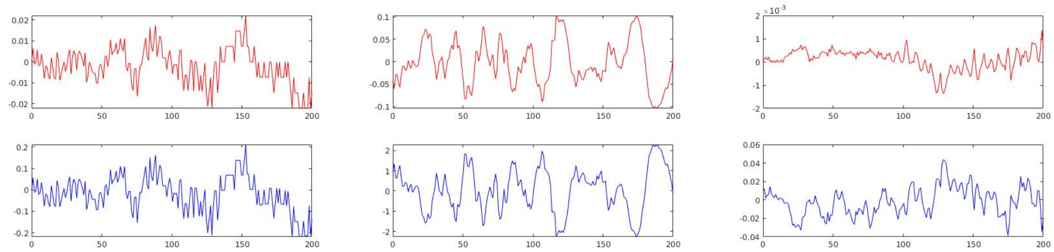
Figure 3: Input (red), Mixed(green), Recovered(blue) Signals

The recovered signals can be in any random order, some were scaled and inverted. To plot them corresponding to their input signal, the values were normalized between -1 and 1. After performing normalization, the input, output and mixed signals have been plotted in Figure 3. It can be seen that for signal 1 and 5, exact reconstruction is achieved, whereas for signals 2,3 and 4, they aren't. To further analyze in detail, only 200 time frames were plotted for the different signals as shown in Figure 4. In Figure 4a, we get an exact mapping of the original signal. Whereas, in Figure 4b and 4c, we can observe that the output signal is inverted and is scaled up respectively.

4 Discussions and Conclusion

A maximum accuracy of **0.99** was achieved i.e. we could re-construct the signals almost as similar to the original signal. The audio was also played and heard and with very little noise, the recovery of the individual source signals was possible.

It has been observed that as we increase the number of iterations, the accuracy of the model also increases. Mixing was performed with all signals and still the original were recovered.



(a) Reconstruction - exact (b) Reconstruction - inverted (c) Reconstruction - scaled

Figure 4: Reconstruction of signals